School of Engineering & Technology

Department of Artificial intelligence and Data Science

AY 2024-25

A

PBL WORK

"Healthcare Data Analysis and Management System Using Power BI"

Submitted to the D. Y. Patil University

**In partial fulfilment of the requirements of the degree of
Bachelor of Technology**

(Artificial intelligence and Data Science)

By

Name of Student: Jadhav Ajay Ganpati

Roll No: A-38

Department of Artificial intelligence and Data Science

CERTIFICATE

This is to certify that **PBL** work entitled

"Healthcare Data Analysis and Management System Using Power BI"

Submitted by

Name of Student: Jadhav Ajay Ganpati

have completed all the Term Work & Practical Work in the subject **Project Based Learning (PBL)** satisfactorily in the department of Third Year Engineering as prescribed by **D. Y. Patil University, Ambi, Pune**, in the academic year 2024 -2025

Prof. Madhavi Patil   **Dr. Vivek Patil**

(PBL Coordinator)   (HOD CE/IT)

ABSTRACT

In the modern era of data-driven decision-making, the healthcare industry stands to gain significantly from effective data analysis and visualization techniques. This Project-Based Learning (PBL) work presents the design and development of a Healthcare Data Analysis and Management System using Microsoft Power BI. The project utilizes a structured healthcare dataset comprising various interrelated entities such as patients, diagnoses, procedures, visits, insurance, departments, and providers.

The core objective of this project is to derive meaningful insights from healthcare data using the powerful features of Power BI. These include real-time dashboards, interactive visualizations, and data modeling capabilities. The project began with importing and transforming data from Excel sheets into Power BI, followed by cleaning, relationship mapping, and creating calculated columns and measures using DAX (Data Analysis Expressions).

Interactive dashboards were built to visualize critical metrics such as the most common diagnoses, department-wise patient distribution, procedure frequency, provider workloads, and insurance utilization. These visualizations empower healthcare administrators and stakeholders to make informed decisions quickly and efficiently.

Using Power BI significantly enhances the analytical depth and interactivity compared to traditional tools like Excel. With automated refresh options and drill-through features, users can explore data from multiple perspectives, leading to faster and more accurate insights.

Through this project, students not only gained practical experience with real-world datasets but also acquired proficiency in Power BI, an industry-standard tool for business intelligence. management, and strategic planning.

TABLE OF CONTENTS

Chapter 1: Synopsis

The healthcare industry is one of the largest producers of data globally. To better serve patients and operate efficiently, healthcare providers must harness this data. In this PBL project, we explore an Excel-based healthcare dataset containing various dimensions such as patients, visits, providers, diagnoses, and more.

The healthcare industry produces massive volumes of data across various departments and services. Managing and extracting insights from such data is crucial for efficient healthcare delivery and administration. This Project-Based Learning (PBL) initiative is focused on developing a comprehensive Healthcare Data Analysis and Management System using Microsoft Power BI.

The dataset used in this project includes multiple Excel sheets representing real-world healthcare entities such as patients, diagnoses, procedures, visits, providers, departments, and insurance records. The data was imported into Power BI where it underwent a series of transformation steps including data cleaning, relationship mapping, and normalization.

Power BI's powerful data modeling and visualization capabilities were used to build insightful dashboards and reports. These included visual representations of the most frequent diagnoses, top-performing departments, insurance usage patterns, patient visit trends, and provider-specific performance metrics.

This project enhances understanding of real-world data analytics and demonstrates how business intelligence tools like Power BI can transform healthcare data into strategic assets. The system developed in this project is scalable and can be extended to include predictive analytics in future versions.

Objectives of the Project:

- To analyze and manage healthcare data using Power BI.

- To identify trends and patterns in healthcare services.

- To develop interactive dashboards for real-time decision-making.

- To gain hands-on experience with Power BI tools, data modeling, and DAX functions.

- To understand the impact of data visualization in the healthcare industry.

- To create calculated fields and measures using DAX (Data Analysis Expressions) for advanced analysis.

- To utilize slicers and filters for dynamic, user-driven data exploration.

- To identify relationships between various healthcare dimensions (e.g., providers vs. departments, visits vs. diagnoses).

- To monitor patient flow and department load for resource optimization.

- To evaluate insurance coverage trends and their impact on hospital revenue.

- To practice real-world BI report development aligned with industry standards.

Chapter 2: Introduction

In recent years, the healthcare industry has seen a massive shift toward digital record-keeping and data-driven decision-making. Hospitals, clinics, and health insurance providers generate and collect vast amounts of data on a daily basis—from patient demographics and diagnoses to insurance claims and provider details. However, raw data alone is not enough. The real value lies in transforming this data into actionable insights that improve efficiency, enhance patient care, and support strategic planning.

This Project-Based Learning (PBL) report focuses on developing a Healthcare Data Analysis and Management System using Microsoft Power BI, one of the most widely used business intelligence platforms today. The project utilizes an Excel-based dataset composed of multiple interrelated entities, including patients, visits, diagnoses, departments, procedures, insurance, and providers. These datasets were loaded into Power BI, cleaned, and linked through defined relationships to build a robust data model.

Using Power BI, interactive dashboards were created to visualize critical healthcare metrics such as patient volume, diagnosis frequency, department workload, provider performance, and insurance coverage distribution. These dashboards are dynamic, enabling users to filter, drill down, and explore the data from various perspectives.

The main purpose of this project is to demonstrate how healthcare data can be transformed into valuable insights using modern tools like Power BI. This not only helps hospital administrators and policymakers make informed decisions but also showcases how data analytics can be applied to improve real-world healthcare systems.

By the end of the project, students developed a deeper understanding of data analytics, data modeling, and visualization, while also gaining practical experience in using Power BI to solve real-life problems.

- o Problem Statement:

Currently, hospital administrators and clinical teams are burdened with manually generated monthly or quarterly reports which often contain outdated information and lack interactivity.

As a result, trends and anomalies such as rising readmission rates, bottlenecks in patient flow, or underutilized resources remain undetected until they escalate into significant issues.

Moreover, decision-makers are forced to work with generalized summaries rather than granular, department wise insights, hampering their ability to take proactive steps. This results in reactive rather than proactive management.

o This project aims to design and implement a dynamic dashboard solution tailored for healthcare institutions, with the following scope and objectives:

• Data Integration: Collect and unify data from EMRs, Patient Administration Systems, and financial records into a single analytical model.

• ETL Pipeline Design: Cleanse, transform, and load data using Power Query to ensure consistency, accuracy, and reliability.

• Data Modelling: Establish logical relationships using star schema design, allowing efficient data slicing and cross-filtering.

• KPI Development: Define and calculate essential healthcare metrics using DAX functions such as ALOS, Bed Occupancy Rate, Readmission Rate, etc.

• Interactive Visualization: Create user-friendly dashboards with capabilities like slicers, time filters, tooltips, and drill-through options for granular analysis.

• Deployment & Testing: Publish the dashboard on Power BI Service, ensuring accessibility for different stakeholders, and refine based on iterative feedback.

Chapter 3: METHODOLOGY

The methodology adopted for this project follows a systematic approach, starting from data acquisition to the final development of interactive dashboards in Power BI. The focus was on transforming raw healthcare data into actionable insights that can support decision-making and strategic planning. Below is a step-by-step explanation of the process followed:

1. Data Collection: The dataset used in this project was provided in the form of multiple Excel sheets, each representing different aspects of a healthcare system. The main tables included:

- Patients

- Providers

- Visits

- Diagnoses

- Departments

- Insurance

- Procedures

- Cities

These datasets collectively simulate real-world healthcare scenarios.

2. Data Preparation and Cleaning: Once imported into Power BI, the data was cleaned using Power Query Editor:

- o Removed null and duplicate entries.

- o Standardized date and number formats.

- o Split and merged columns where necessary.

- o Renamed fields for better readability.

Data relationships were defined among tables (e.g., linking Patients with Visits, Diagnoses with Procedures) to create a normalized model.

3. Data Modeling: The cleaned tables were linked using one-to-many and many-to-one relationships. A star schema approach was used to optimize performance and maintain clarity. Key features:

Primary and foreign keys identified across datasets.

Created relationships to allow cross-filtering and drill-downs.

4. Creating Calculated Columns and Measures: Using DAX (Data Analysis Expressions), several custom calculations were added:

- o Total number of patients per department.

- o Most frequently diagnosed conditions.

- o Insurance usage by patient category.

- o Average visits per provider.

These metrics helped in extracting valuable insights from the data.

5. Data Visualization and Dashboard Design: Multiple dashboards and reports were created in Power BI using:

Bar charts, Pie charts, and Line graphs

Tree maps, Card visuals, and Slicers

Drill-down functionality and tooltips

Dashboards were made user-friendly and interactive, allowing stakeholders to explore data with ease.
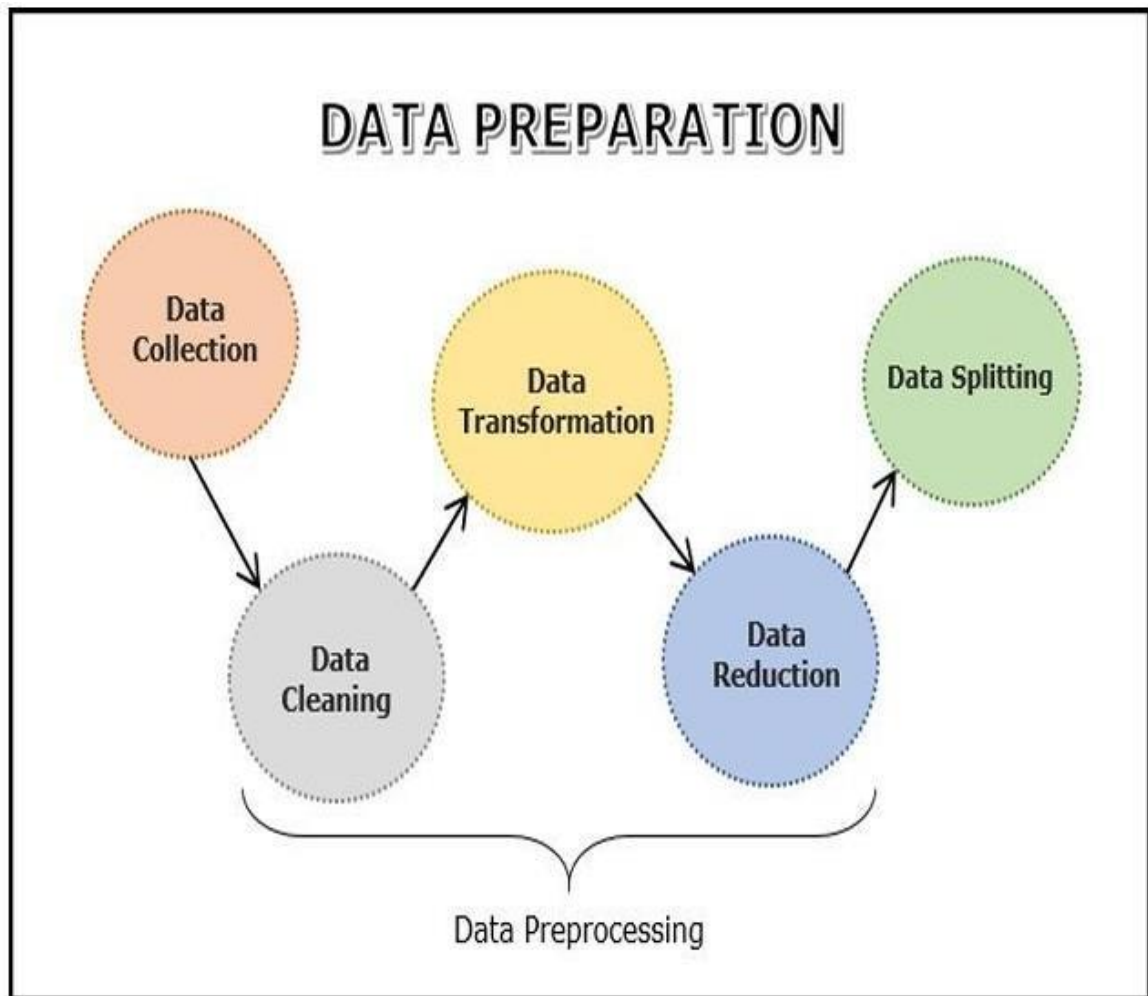
6. Analysis and Interpretation: Visualizations were interpreted to extract business insights such as:

- o Departments with highest patient inflow

- o Provider workload distribution

- o Popular procedures

- o Diagnosis trends over time

These findings can be useful for hospital management and policy-makers.

7. Final Review and Optimization: The dashboards were tested for performance and usability. Filters and slicers were added for end-user control. Report themes and layouts were finalized for a professional look and feel.

Data Preparation:

DATA PREPARATION

Data Collection

Data Cleaning

Data Transformation

Data Reduction

Data Splitting

Data Preprocessing

Data Preparation

• Data Cleansing: Initial preprocessing involved handling missing entries, duplicates, and anomalies. Inconsistent date formats were standardized, and erroneous values (e.g., negative patient stays) were identified and corrected.

• Normalization: Codes and identifiers for diagnoses, departments, and patient types were normalized using mapping tables to ensure uniformity.

• Dimension Table Creation: Lookup tables were constructed for entities such as Time, Department, Diagnosis Code, and Gender to support efficient filtering and relational integrity.

• Data Transformation: Advanced transformations included merging datasets on common identifiers, creating calculated columns for metrics like Age, and segmenting patient groups based on length of stay or condition severity.

Chapter 4: ADVANTAGES, DISADVANTAGES AND APPLICATIONS

- ## *Advantages of Using Power BI in Healthcare Analytics:*

Interactive Visualizations:

Power BI enables the creation of dynamic dashboards that help healthcare professionals quickly interpret patient and treatment data.

Easy Integration:

Easily connects to Excel files like your patients, visits, and procedures datasets without the need for complex ETL processes.

Real-time Data Monitoring:

Power BI allows for real-time updates which is crucial for tracking patient visits, diagnoses, and insurance claims.

User-Friendly Interface:

Drag-and-drop features make it accessible for non-technical users such as healthcare staff or administrators.

Customizable Reports:

Reports can be tailored for departments, doctors, or patient-specific dashboards for better decision-making.

- *<u>Disadvantages of Using Power BI:</u>*

Limited Data Cleaning Capabilities:

Complex data transformation still requires Power Query or external tools, which may be less intuitive.

Performance Issues with Large Data Sets:

Handling very large healthcare datasets might slow down the dashboard performance without optimization.

Requires Licensing for Full Features:

Some features (like sharing dashboards with others securely) require Power BI Pro or Premium licensing.

Data Security Concerns:

In healthcare, handling sensitive patient data requires strict compliance (HIPAA, etc.), and cloud-based tools may pose concerns if not handled properly.

- *<u>Applications of Power BI in Healthcare Analytics:</u>*

Patient Demographics & Trends:

Analyze patient data from the patients and visits sheets to understand age distribution, common diagnoses, or location-based health trends.

Department Performance Tracking:

Use the departments, procedures, and providers sheets to assess departmental efficiency and workload.

Insurance Claim Analysis:

Visualize the number of claims, types of insurance used, and approval rates using the insurance data.

Diagnosis & Procedure Correlation:

Cross-reference diagnoses and procedures to find common treatment patterns and success rates.

Resource Utilization:

Track how often different departments or procedures are used, helping in better hospital resource management.

Chapter 5: CONCLUSION

This Project-Based Learning (PBL) work demonstrated the practical application of Microsoft Power BI in analyzing and managing real-world healthcare data. The project successfully highlighted how structured data, when modeled and visualized effectively, can lead to actionable insights that support operational and strategic decision-making in the healthcare sector.

By using Power BI, we were able to create dynamic dashboards that visualized key metrics such as patient distribution, provider performance, diagnosis trends, and insurance utilization. The data was cleaned, modeled, and enhanced with calculated measures using DAX, enabling the creation of professional-level reports and dashboards.

Throughout this project, we gained valuable experience in:

- o Importing and transforming data from multiple sources.

- o Building relationships between interdependent datasets.

- o Designing interactive dashboards that provide real-time insights.

- o Applying analytical thinking to solve real-world problems in healthcare.

In conclusion, this project not only strengthened our technical skills in data analysis and visualization but also provided insight into how healthcare systems can benefit from business intelligence tools like Power BI. It lays a strong foundation for future projects involving predictive analytics, machine learning, and AI applications in health informatics.

➢ **Dashboard for Healthcare Analysis Dashboard**:
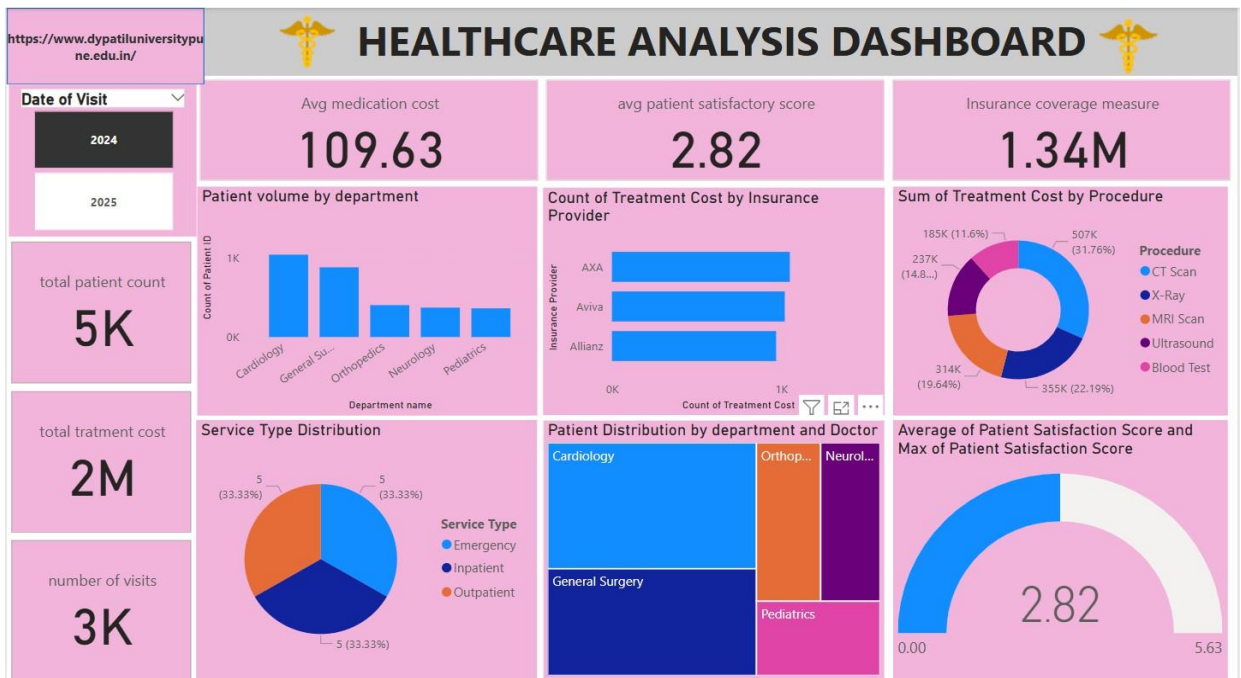
Chapter 6: REFERENCES

- o Microsoft Power BI Documentation

[Retrieved from https://docs.microsoft.com/en-us/power-bi/ - Search](https://docs.microsoft.com/en-us/power-bi/)

- o Sample Healthcare Dataset (Excel Files)

(Used for analysis of patient records, visits, diagnoses, and procedures.)

- o D. Y. Patil University Guidelines

(For report structure and formatting.)

- o Articles on Healthcare Analytics

Example:

Kankanhalli, A., et al. (2016). Big data and analytics in healthcare: Introduction to the special section. Information Systems Frontiers.

- o YouTube tutorials on Power BI dashboards for healthcare

(Used for understanding visualization techniques.)